Carbon Dioxide Emissions, economic growth and energy mix: empirical evidence from 93 countries

Elena Stolyarova[1]

This paper provides an empirical analysis of CO2 emissions and economic growth, based on a panel dataset covering 93 countries over the period 1960-2008, and examines the challenge of country selection for homogenous and appropriates groups. Foremost, we have proposed a no parametric hierarchic clustering approach, based on 17 criteria and used the Hierarchical Clustering on Principal Components (HCPC). The results of the clustering indicate the optimal partition of 93 countries into 7 groups, each with its own characteristics. The unit root and cointegration tests show that the long-run relationship does not exist for any clusters and the nature of stationary is different between and into the groups. This result indicates the short-run relationship between CO2 emissions and its determinants. Dynamic Panel Data and WITHIN models were estimated to explain the growth rate of per capita CO2 emissions. It's found that the growth rate of per capita CO2 emissions depends positively on the growth rate of per capita GDP and negatively on the growth rate of energy mix.

Keywords: CO_2 emissions, Panel Data Analysis, Economic Growth, Hierarchical Clustering on Principal Components.

1. Climate Economics Chair,
elena.stolyarova@chaireeconomieduclimat.org

Introduction

Between 1960 and 2008, emissions of CO_2 from fuel combustion have tripled and the main actors have changed. In 1960 the contribution of emissions by China was around 9%, 1% for India and 10% for rest of the world. By 2008, their contribution was 24%, 5% and 23% respectively, and China becomes the largest emitter in the world. This growth is the results of 210% increase on energy consumption, much of which was met by a 370% increase in electricity generation. The share of electricity's emissions in the total emissions of CO_2 was between 18% and 78% in 2008 (WDI, 2012). The variation between the countries can be explained by many determinants among which are the level of income and the energy mix.

Most previous studies of CO_2 – Income relationship aim either to verify and estimate the Environmental Kuznets Curve (EKC) hypothesis of economic inequity or to describe the long-run equilibrium relationship between GHG emissions and energy consumption, or GDP, or other. The EKC bases on hypothesis of U-inverted relationship between income level of country and level of GHG emissions. In low income countries energy use and GHG emissions increase as income increases. As income levels increase, societies have the awareness and means to implement costly environmental schemes, leading to reductions in emissions. So the level of emissions decreases from a certain level of income. The first application of Kuznets Curve to environmental studies is done by Grossman and Krueger (1991, 1993, 1995) [1], who were followed by Holtz-Eakin (1995) [2], or more recently by Perman and Stern (2003) [3], McKitrick and Strazicich (2005) [4], Aldy (2006) [5] and Dinda (2004) [6]. The results of these studies are controversial about EKC's hypothesis, giving opposite conclusions.

Studies of long-run equilibrium relationships, which are often complementary to the EKC studies, focus on long-run causal relationships between several variables affecting the level of emissions or energy consumption. Energy consumption and economic growth are commonly the focus. For example, Ang (2007) [7] proposed a Vector Error Correction Model for France, Narayan and Smith (2008) [8] studied the Energy Consumption-Economic growth nexus for the G7 countries. Mehara (2007) [9] investigated the same relationship in the case of oil producing countries. Finally, Halicioglu's (2009) [10] studied on the Turkish long-run elasticity between energy consumption and economic growth. There has been relatively little work published on the relationship between GDP and emissions growth, but some examples can be done. Dinda and Coondoo (2006) [11] performed cointegration analysis between per capita CO_2 emissions and per capita GDP on a panel of 88 countries and conclude that a long-run relationship exists between the variables. Lee and Lee (2009) [12] applied a different methodology - the more efficient panel unit root test (SURADF) and concluded differently: in several regions, the long-run relationship does not exist. Marrero (2010) [13] proposed a dynamic panel data model that combines the EKC hypothesis for nexus GHG emissions-Income and energy mix for EU27. He concludes that European countries converge on per capita GHG emissions while EKC hypothesis are rejected. Hossain (2011) [14] examined the short-run and long-run relationship between per capita CO_2 emissions, energy consumption, per capita GDP, trade openness and urbanization for NIC countries, and found that the long-run elasticity is higher than short-run elasticity.

The econometric approach, which is usually used to estimate the relationship between GHG emissions and economic growth, as well as to test EKC hypothesis, has been criticized in academic literature on many points. Dijkgraaf and Vollebergh (2005) [21] tried to verify the assumption of the same emissions-income elasticity across countries (regions) and time. It is a strong assumption that is in generally difficult to observe. The authors found that the turning point is different for each OECD country and that a slight change in the model leads to a significant change in the estimated parameters. Aslanidis (2009) [22] proposes a critical survey of the econometric techniques used to estimate EKC hypothesis. These techniques refer to baseline model, homogeneity across countries, fallacious test of unit root and cointegration, spurious regressions and spatial dependence. M. Wagner (2008) [23] also criticizes econometric approach used to estimate EKC. He focuses on studying the problems of non-linear transformation and unit root tests. These critics led us to leave the model with EKC hypothesis and to find a solution of weakness Panel Data Model.

Some studies attempt to refine the models by assembling the countries in homogeneous groups. The most popular techniques of building the sub-panel are based on group selection, according to either one qualitative or quantitative variable. Sharma (2011) [15] groups 69 countries using income level. Niu (2011) [16], Dinda (2006) [11], Lee (2009) [12] and Narayan (2010) [17] assemble the countries by region. Wang (2012) [18] analyses the relationship between CO2 emissions from oil and economic growth by pooling according to economic growth regime. Some studies focus on groups of economic union or with the same level of economic development - for example, Hossain (2011) [14] for Newly Industrialized Countries, Pao (2010) [19] for BRIC, Lee (2008) [20] for OECD countries. Mehrara (2007) [9] works on a panel of 11 oil exporting countries.

However, groupings by country can make some problems for panel data analysis approach. Panel analysis is used to enhance econometric estimations and to obtain more robust results thanks to their two-dimensional data: time and individual. Full or partial homogeneity in parameters is necessary to estimate the model and so the appropriate selection of individuals to compose a panel is very important. Grouping based only on the region division or on the economic level could recover inadequate and lead to a very imprecise estimation of parameters, as well as weakening the power of the tests. In a case of panel data analysis of GHG's determinants, several factors must be considered other than mentioned above, like reserves of fossil fuel, energy mix, size of country, energy and environmental policies, level of production, level of trade and many other. The countries with the same level of economic development may have different relationship between emissions and economic growth for many reasons. For example, small island countries generate electricity from oil, because this solution is the least expensive, and so they have high level of emissions per capita. Some developing countries can have more Renewable Energy Sources (RES) than developed countries and so they have a lower emissions level. Energy mix can be different and depends on domestic resource endowments.

The purpose of this paper is twofold: first, to investigate the short-run nexus between CO2 emissions, income and energy mix, and second, to propose homogeneous groups of countries in order to obtain a more efficient model. To confirm the choice of a short-run model, we are use cointegration analysis to demonstrate the absence of cointegration between CO2 emissions and income. We propose a WITHIN model with an evolving mix of alternative and nuclear energy use.

The rest of this paper is organized as follows: Section 1 offers some reflections on the question on the issue of choice determinants. Section 2 provides a brief survey of studies on countries' homogeneity. The data and some descriptive statistics are presented in section 3. Section 4 presents the methodology and the results of clustering to make groups of homogenous countries. Section 5 describes the econometric approach for modeling the short-run relationship between CO2 emissions, economic growth and energy mix, and the results of model specification tests. Section 6 presents the model and its results. At least, section 7 concludes with final remarks.

1. Determinants of CO_2 emissions

Carbon Dioxide emissions have two principal sources. The largest source is the combustion of fossil fuels, such as natural gas, crude oil and coal. The second source is from industrial processes that emit CO2 as a result of a chemical reaction. In both cases, CO2 emissions – energy consumption nexus is mainly chemical and physical. But the relationship fluctuates according to the types of combustion, means of production and energy generation, energy efficiency etc. This suggests that the two relationships CO2 emissions-GDP and Energy consumption-GDP do not necessarily have the same elasticity or follow the same trend over time, in which case energy consumption can only partially explain the variation of CO2 emissions. Additionally if energy consumption and CO2 emissions are coupled in a linear relationship, we cannot use these variables in the same model for reasons of collinearity.

The literature on the relationship between energy consumption and economic growth is rich in examples of application and provides a unanimous conclusion on the existence of the relationship between energy consumption and economic growth. By contrast, the literature on the emissions-economic growth nexus is scarce. Dinda and Coondoo (2004, 2006) [6], [11] study a panel of 88 countries grouped by region to examine the long-run nexus between per capita CO2 emissions and per capita GDP by using IPS panel root test and Engel-Granger methodology of cointegration. The results show that series have a same order of integration and that a long-run relationship exists for Africa, Central America and Europe. Lee and Lee (2009) [12] devote their paper to study the order of integration of per capita CO2 emissions and per capita GDP, using SURADF unit root test, to consider the presence of cross-country correlation in the data. They conclude that two variables don't have the same order of integration and so the long-run relationship between emissions and income does not exist. Naryan and Naryan (2010) [17] examined the EKC hypothesis for 43 developing countries and estimated long-run and short-run elasticity between emissions and income, taking into account the problems of collinearity between income and income-squared, which is usually used as EKC model specification.

This is not the first time that model specification and variable selection are criticized for example, see Perman and Stern (2004) [3]. More recently, Sharma (2011) [15] propose a short-run multivariate model that considers trade openness, urbanization, per capita electric power consumption and per capita total primary energy consumption, all in addition to per capita CO2 emissions and per capita GDP. The author concludes that for all countries, except for those with high income, only per capita GDP and urbanization are determinant for CO2 emissions. But the author uses both GDP and energy consumption as CO2 emissions' determinants. This may distort the results.

To solve this problem, we need that the energy variable which is related to CO2 emissions, does not depend on per capita GDP and explains energy consumption or mix. Alternative and nuclear energy use in % of total energy use (ANEU) is a good specification which meets the requirements. First, ANEU reflects all sectors, including the power sector. Second, an increase of this variable explains the decrease of level per capita CO2 emissions. And finally our tests show that ANEU is not correlated with GDP per capita.

2. Countries' homogeneity in econometric studies

Several authors have studied the problem of homogeneity. Burnside (1996) [24] analyzed a panel of US manufacturing industries and concluded that production function is not homogenous for all industries. Baltagi and Griffin (1997) [25] studied the determinants of gasoline demand in OECD countries and rejected the hypothesis of coefficients' homogeneity. Furthermore, Baltagi, Griffin and Xiong (2000) [26], as well as Baltagi, Bresson and Pirotte (2002) [27] emphasize the degeneration of results when observation number is large. In case of misspecification of the countries' group, not only the estimations are wrong, but all tests are not consistent. Number of solution has been proposed to cluster the countries in homogeneous groups for Panel Data Analysis. Durlauf and Johnson (1995) [28] suggest to detail individuals to groups, using regression tree analysis. Information criterion method was proposed by Kapetanious (2006) [29]. This method is based on a concept of hierarchical clustering, and leaves the individual estimates. But such a practice is questionable due to a small number of observations. V. Sarafidis and N. Weber (2009) [30] proposed a method based on the principle of partitional clustering, which can circle the problem of small number observations. They use hill-climbing algorithm which tries to find a best partition in predetermined number of group. The algorithm is based on minimization of total residuals sum of squares (RSS) obtained for each cluster partition's estimation. But in case of atypical initial partition, the results may conduct to wrong class partition. Moreover this method is only based on model estimation and is sensible to model specification.

In our paper, we propose hierarchical clustering on principal components (HCPC) which presents some advantages. The HCPC does not need any econometric modeling, it uses the resemblances and differences among the individuals from multivariable criteria. It can be applied in a case of non-linear relationship between the series, and does not depend on the number of observations[1]. We can also consider the additional data set of variables, which can explain the different income-emission elasticity between the countries, and at the same time, does not affect the relationship between per capita emissions, per capita GDP and percent of alternative and nuclear energy use. In addition to the variables already used, we use the following variables to groups the countries: the level of CO_2 emission and GDP, country's population, percentage of different sources used to produce electricity, contribution of power and industrial sectors in CO_2 emissions, production of fossil fuel (oil, natural gas and coal), as well as geographical criterion. To solve the nonlinear relationships problem, and eliminate disparities between variables, all quantitative series are transformed to qualitative data set. We consider that qualitative variables are insensitive to the time, so that only last year is considered to perform the classification. For example: if the country does not produce the oil at the moment, it will not become the largest producer next year.

3. Data presentation and descriptive statistics

We used annual unbalanced panel data covering the period from 1960 to 2008 for 93 countries. The data of CO_2 emissions (Mt) was obtained from the Climate Analysis Indicators Tool (CAIT); the data of GDP (in $ at 2000 prices), population and percentage of alternative and nuclear energy use (ANEU) from the World Development Indicators (WDI) of the World Bank. We also used an additional dataset which allowed us to perform hierarchical clustering. Electricity production from oil, coal, RES, nuclear, natural gas and hydraulic resources (in percent of total electricity production); contribution of power and industrial sectors in total CO_2 emissions (in percent of total CO_2 emissions) are provided from the WDI. The production of oil (Mt), coal (Mt) and natural gas (Mteo) derives from BP statistical review of world energy.

The description of panel countries composition is given in **Annex A** and descriptive statistics for principal data set are given in **Table 1**. The 93 countries of our data set cover 97% of worldwide CO_2 emissions. The five countries with the highest mean per capita emission over time are: Luxembourg (33.36), Qatar (30.47), Kuwait (25.04), United Arab Emirates (24.26) and United States (19.57). In 2008, a per capita CO_2 emission was the highest in the Qatar (30.47) and the lowest in the Sudan (0.2). The highest level of per capita CO_2 emissions over time was in Luxembourg in 1961 (55.55) and in Qatar in 2005 (55.30).

[1] Under the assumption that membership in a group does not change over time

The level of per capita GDP varies from 72 to 61374 $ per capita across the time. The mean per capita GDP is highest for: United Arab Emirates, Switzerland, Qatar, Luxembourg and Japan. In 2008, the five richest countries were Luxembourg (55.8)[2], Norway (41.4), Japan (40.2, United States (38.3) and Switzerland (38.2). Nigeria has the lowest level of per capita GDP in 2008 (0.5).

The mean of % of alternative and nuclear energy use varies from 0% to 82.9%. 11 countries of our dataset do not use any clean energy. Iceland (82%), Sweden (45%), France (45%), Norway (40%) and Switzerland (39%) present the best average performance in 2008.

Table 1 : Summary of basic statistics

	CO_2 per capita	GDP per capita	ANEU (in %)
Mean	6.55	8078	7.36
Median	5.05	3666	2.44
Maximum	55.54	61374	82.86
Minimum	0.05	72	0
Std. Dev.	7.05	9609	11.8
Observation	4540	3664	3472
Cross sections	93	93	93

The additional data set is present in **Table 2**. We take CO_2 emissions, GDP and population to consider economic and geographical scale of countries. The percentage of electricity production from different sources allows us to take into account countries' electricity mix. CO_2 emissions from electricity and industrial sectors give an overview of sectorial contribution in CO_2 emissions. Oil, natural gas and coal productions have a double effect on emissions. First the production of fossil fuels is polluting itself, secondly, the energy mix depends on fossil fuel endowments. Finally geographical variable describes advantages and facilities between the countries in same region.

Table 2: Description of additional data set

Variable	Description		Categories
CO2	CO_2 emissions	Mt	
kCO2	Per capita CO_2 emissions	Per capita Mt	
GDP	GDP	In $ at 2000 prices	
kGDP	Per capita GDP	Per capita, In $ at 2000 prices	
POP	Population		4 categories (very low, low, high, very high)
ANEU	Alternative and nuclear energy use	in % of total energy use	
CO2_elec	CO_2 emissions from power sector	In % of total CO_2 emissions	
CO2_indus	CO_2 emissions from industrial sector	In % of total CO_2 emissions	
Elec_coal	Electricity production from coal	in % of total production	
Elec_oil	Electricity production from oil	in % of total production	
Elec_hydro	Electricity production from hydraulic	in % of total production	
Elec_gas	Electricity production from natural gas	in % of total production	
Elec_nucl	Electricity production from nuclear	in % of total production	4 categories (no, low, middle, high)
Elec_RES	Electricity production from other RES	in % of total production	
Prod_oil	Production of oil	Mt	
Prod_gas	Production of natural gas	Mt	
Prod_coal	Production of coal	Mteo	
Region	Geographical variable	Africa, Middle East, Asia and Pacific, Eastern Europe, Western Europe, North America, Central America and South America	

[2] In k$ per capita

4. Classification

The Principal Component Methods are usually used to transform the data in a new set of variables (principal components) done by eigenvalue decomposition before applying the hierarchical clustering. The procedure depends on the type of variable and on how they relate to each other. In the case of quantitative variable and linear relationship between the variable, Principal Component Analysis (PCA) is used. Multiple Correspondence Analysis (MCA) is applied for categorical variables and based on χ^2 distance. It is possible to perform principal component method for mixed data. One of the solutions was given by Hill and Smith (1976) [31].

We chose to transform a data from quantitative to categorical variable and apply MCA procedure for many reasons. First, several variables have outliers: they may not be ignored and cannot be used without distorting the results of clustering. For example, the biggest countries (China, USA, Canada, Russia, India etc.) have a very high level of population, income, CO_2 emissions and they extract a lot of fossil fuel. After transformation on principal components variables and PCA application, they will have very high coordinates and will be far from center of mass, as well as far from other countries. As a result, HCPC based on PCA application rank atypical countries at least and they will have their own cluster. Secondly, some variables are not linearly associated as the percentage of nuclear and RES power production, or as the natural resource endowment does not depend on the level of GDP. Finally, the regional partition is important: economic ties between the countries, access to the same resources and infrastructure, etc.

To perform the HCPC, we use R package for multivariate analysis (FactoMineR) developed by Lê, Josse and Husson (2010) [32]. The clustering procedure is based on Ward's criterion which allows us to choose the pair of clustering to merge at each step by minimizing the growth of within inertia, i.e. minimization of the reduction of between inertia. Optimal number of cluster can be chosen by comparison with the level of total inertia. In general the optimal number of cluster is two or three, but in our case we need a finest class composition in a range of 5-10 clusters.

The class composition and their characteristics are present in **Table 3**. The geographical countries' distribution in groups is shown in **Figure 1**. The countries are distributed in 7 clusters according to 17 criteria. The electricity mix and geographical criteria characterize most clusters, as well as fossil fuel endowments. Only cluster 1 is specified by the economic size of countries (very high level of CO_2 emissions and GDP). The cluster 4 is characterized exclusively by medium level of nuclear and RES electricity production. It may be noted that the EU countries are shared out among four classes: rich countries in Western Europe with a lot of clean energy; countries from Eastern Europe with high level of coal power source; the countries with a balance electricity mix and finally United Kingdom belongs to the same cluster as USA, China, Russia etc. This result is interesting because it gives an unusual distribution which has never been applied. In the next section we will verify if our classification allows having groups of homogeneous countries.

5. Econometric approach

In the case of insufficient data observations and the presence of many same individuals, the panel data analysis is the most appropriate method to estimate econometric models, thanks to its double temporal and individual dimensions. Assembling several individuals leads to enhance an econometric model. To confirm the validity of the final model, it's necessary to apply different tests. First, we need to know if grouping individuals makes sense. In the case of group's misspecification, the rest of the tests and the econometric model can lead us to the wrong conclusion. So, begin with test parameters' homogeneity seems important. After that, the unit root and cointegration tests should be performed to determine the nature of stationary and justify the choice of a short-term model. Finally the dynamic regression model is estimated to understand the connection between the series.

5.1 To Pool or not to pool?

Panel Analysis is based on hypothesis of coefficients' partial homogeneity for selected groups of individuals. But two extreme cases exist: Either the relationship between studied variables is different for each individual, or the individuals are perfectly homogenous. In the first case panel analysis makes no sense and the individual models should be estimated. In the other case, a simple pooling model is enough. We take an interest in the test procedure proposed by Hsiao (2003) [33], which is simple; apply in the case of a model with fixed effects. He suggests 3 consecutive tests. Foremost he tests under null hypothesis perfect homogeneity of parameters against full heterogeneity. After that, he tests under null hypothesis partial homogeneity of parameters (constant slopes and different intercepts for each country) against full heterogeneity. And finally partial homogeneity is tested against perfect homogeneity of parameters. The test's statistics follow Fisher's distribution. A slight modification is necessary to take into account time fixed effects. Given that the number of observations equals the number of time dummy, we cannot estimate for each country the model with 49 time fixed effects for hypothesis of full heterogeneity. Nine dummies are introduced to circle this problem. They represent the first and the second oil crises, Asian financial crisis and the last global financial crisis. We propose three models:

$$CO2_{n,t} = \alpha + \gamma * GDP_{n,t} + \varepsilon_{n,t} \quad (1)$$

$$CO2_{n,t} = \alpha_n + \beta_t + \gamma * GDP_{n,t} + \varepsilon_{n,t} \quad (2)$$

$$CO2_{n,t} = \alpha_n + \sum_{i=1}^{9} d_i + \gamma_n * GDP_{n,t} + \varepsilon_{n,t} \quad (3)$$

Table 2 : Composition of the groups and their characteristics

Class	Class composition[3]	Characteristics of classes (for majority of countries)	Representative Country
N° 1	AUS, CAN, CHN, ***GBR***, IDN, IND, MEX, RUS, THA, TUR, USA	- Countries with very high level of CO2 emissions and GDP (no per capita, in volume); - Producers of fossil fuel.	Russia, India, China
N° 2	***AUT, BEL***, CHE, ***DNK, FRA,*** GTM, ***IRL***, ISL, ***ITA, LUX, NLD, NOR,*** PHL, ***PRT, SWE***	- Countries from Western Europe; - High level of GDP per capita; - A lot of RES[4] power; - No producer of coal or natural gas.	Suede, Switzerland
N° 3	***BGR,*** BIH, BLR, ***CYP, CZE, EST, GRC,*** HRV, ***HUN,*** JAM, ***LTU, LVA, MLT,*** MNG***, POL, ROM, SVK, SVN,*** UKR, ZAF	- Countries from Eastern Europe; - Rather high level of GDP per capita; - No producer of fossil fuel (oil, natural gas, coal); - Many coal power plants.	Estonia
N° 4	***DEU, ESP, FIN***, JPN, KOR, NZL	- Medium level of RES power; - Medium level of nuclear power.	Japan, Germany
N° 5	AGO, AZE, BGD, CMR, CUB, DZA, EGY, HND, IRQ, JOR, KAZ, MAR, MYS, NGA, PAK, SDN, SYR, TUN, UZB, VNM, YEM	- Countries from Africa; - Low CO_2 emissions per capita; - Low GDP per capita; - Without RES, - Very low % of alternative and nuclear energy.	Angola, Sudan
N° 6	ARE, BHR, IRN, ISR, KWT, LBY, OMN, QAT, SAU, SGP, TKM, TTO	- Middle East countries; - Low % of alternative and nuclear energy; - High level of CO_2 emissions per capita; - Use natural gas for the electric power production.	Qatar, Oman
N° 7	ARG, BOL, BRA, CHL, COL, ECU, PER, VEN	- Countries from South America; - A lot of hydraulic power; - Producers of natural gas (medium level).	Peru, Colombia

[3] The members of EU ETS in bold italic
[4] Renewable Energy Sources

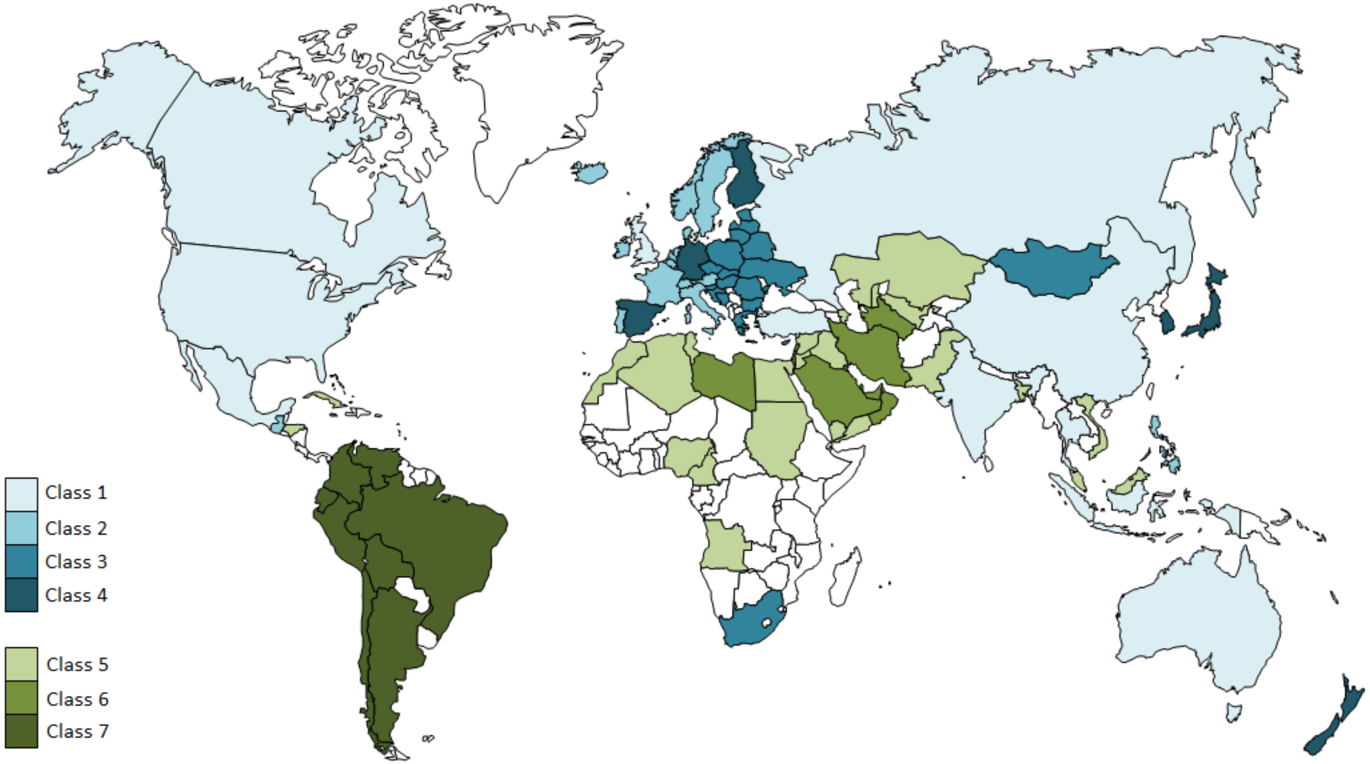

Figure 1 : Geographical distribution of country's sample

Where each equation assumes a relationship between the rate of per capita CO_2 emissions ($CO2_{n,t}$) and the rate of per capita GDP ($GDP_{n,t}$), observed for an individual n at the time t[5]. In the model (1) α is an intercept and γ is an invariant elasticity between emissions and income. In the model (2) α_n describes time-invariant individual fixed effects and β_t describes individual-invariant time fixed effect, γ represents an invariant elasticity between emissions and income. Finally, the model (3) presents N equations for each country where α_n country intercept, d_i – nine time dummies and γ_n - N slopes own for each country. The error term is denoted by $\varepsilon_{n,t}$. Variables are represented by growth rates to make them stationary in the possible presence of the unit root. The first equation presents a perfect homogeneity of parameters. The equation (2) – presents partial homogeneity with the elasticity γ identical for each individual, and finally equation (3) presents full heterogeneity of the parameters. Foremost, we test perfect homogeneity of parameters (1) against at least one of the parameter which is not common for individuals. In the event of rejection of parameter's perfect homogeneity, the test of partial homogeneity (2) is performed. If partial homogeneity is rejected again, panel data analysis cannot be applied and the models for each individual should be estimated; otherwise the grouping of individuals is correct and panel analysis could be performed. In this case, it is possible to realize a third test to confirm the first one, which verifies the homogeneity of parameters outside the own constant terms assumed for each individuals.

The **Table 3** summarizes the results of three tests. The first test allows us to conclude that null hypothesis of perfect homogeneity is rejected for clusters 1, 2, 3, 4 and 7. For cluster 5 and 6 (African and Middle East countries, producers of fossils fuels) the first test suggests that the pooled model is appropriate. The results of a partial homogeneity test suggests that elasticity between the rate of per capita CO_2 emissions and the rate of per capita GDP is the same for all countries, except cluster 1 which represents the largest countries like US, Canada, China etc. For cluster 2 (Western countries with high level of per capita GDP) and 7 (South American countries with a lot of hydraulic power) the null hypothesis of partial homogeneity is not rejected at 1% level. The third test reinforces the choice of a partial homogeneity model with country and time fixed effects, because the results of the third test suggest that time and country fixed effects are not homogeneous. We conclude that panel data analysis can be applied for all clusters and the results of estimations and tests will not be biased, except for the first cluster.

Table 3 : Results of homogeneity tests, F statistic

Cluster	Pooled[a]	Partial[b]	Pooled[c]
Class 1	7.56***	8.12***	1.93***
Class 2	7.43***	2.31**	3.66***
Class 3	2.44***	0	2.91***
Class 4	7.37***	0	2.59***
Class 5	0	0	1.36***
Class 6	1.08	0	3.35***
Class 7	4.42***	2.06**	1.34**

Note:
a – H_0: α, β, γ are invariant against at least one of parameter depending on time or country
b – H_0: model with time and country fixed effects against full heterogeneity of parameters
c – H_0: α and γ is invariant under condition invariant slope against partial homogeneity
d – ***, ** and * indicate statistical H_0 rejection at 1%, 5% or 10% level respectively

[5] We do not interesting in percentage of alternative and nuclear energy use because null value for some countries and we suppose that the main determinant of per capita CO2 emissions is per capita GDP.

5.2 Panel unit roots test

In most cases, the time series data have a non-stationary nature. Several tests of unit roots exist for panel data. We can refer to the test of Levin-Lin-Chu (2002) [34], which bases on the assumption of a common level of integration for all individuals. The weakness of this test lies in the formulation of the test hypotheses: all series have unit root against all series are stationary. Maddala and Wu (1999) [35] have proposed the test with heterogeneous specifications of unit root whith an individual unit root process. A similar test was proposed by Im, Persan and Schin (2003) [36]. Meanwhile, Hadri (2000) [37] suggested a test with opposite hypotheses: stationary of panel data against a presence of unit root. We have chosen to apply IPS test to check the order of integration of study series. The authors specify a separate regression for each cross section:

$$\Delta y_{i,t} = \alpha * y_{i,t} + \sum_{j=1}^{p_i} \beta_{ij} * \Delta y_{it-j} + X'_{it} * \delta + \epsilon_{it}$$

With hypothesis:

$$H_0: \alpha_i = 0, \forall\, i$$

$$H_1 \begin{cases} \alpha_i = 0, & \forall\, i = 1, \dots, N_1 \\ \alpha_i < 0, & \forall\, i = N_1 + 1, N_1 + 2, \dots, N \end{cases}$$

Where presence of unit root is tested against some cross-section have a unit root. The authors propose two statistics: Z-bar with expected value and variance of asymptotic distribution and W-bar with stronger statistical power, especially for a small number of observations and residuals' autocorrelation. They also emphasize that the results of the test without trend are more powerful. To perform this test we apply logarithmic transformation for per capita CO_2 emissions, per capita GDP and % of alternative and nuclear energy use. The results are given in **Table 4**.

Per capita CO_2 emissions with logarithmic filter have a unit root for cluster 1, 5 and 7, either with both trend-intercept or only with an intercept model, whereas cluster 2, 3 and 4 have opposite test's conclusions. For these groups, null hypothesis of unit root are not rejected in the case of a model with trend, and are rejected in the case of a model with only an intercept. This means that several cross-sections do not have a unit root. In the case of per capita GDP cluster 1, 5, 6 and 7 have a unit root. For cluster 2 and 4 the null hypothesis is rejected in test with intercept, for cluster 3 in the case of a test whit trend and intercept. Our results join the conclusions of Lee and Lee (2009) [12]: per capita CO_2 emissions and per capita GDP is stationary for a part of the countries. Only clusters 1, 5 and 7 have a same order of integration for emissions and income and so they may have a long-run relationship between emissions and GDP.

The variable ANEU is not stationary for cluster 3, 4 and 6. This result is interesting: ANEU is measured in percentage and we might have expected stationary of the variable for all clusters. But countries from these groups have to make progress to improve the carbon intensity and to use more clean energy sources. For class 2, the null hypothesis of unit root is rejected; it is probably due to their very high level of electricity production from clean sources. For France, Sweden, Switzerland, Norway and Iceland, electricity production from fossil fuel does not exceed 10 percent. Considering that it is more difficult to use alternative energy for other than electricity sector, the value of ANEU does not increase. The null hypothesis is also rejected for class 5 because of the absence of alternative energy use or a very low value of ANEU. Members of this cluster are fossil fuel producers and low GDP per capita. At the moment they are not interested in improving their performances. For cluster 1 the null hypothesis of unit root in ANEU is rejected in the case of a test without trend, but only India's ANEU is stationary, while for Russia and United Kingdom, ANEU has a trend. So we can conclude that the results from the test with trend and intercept are most probable.

In conclusion, we can tell that the long-run relationship is possible only for cluster 1, between per capita CO_2 emissions, per capita GDP and % of alternative and nuclear energy use. For other clusters, three data series do not have the same order of integration. The cointegration is possible between per capita CO_2 emissions and per capita GDP for cluster 1, 5 and 7.

Table 4 : Panel IPS test for per capita CO2 emissions, per capita GDP and ANEU

W – bar	Log CO_2 per capita		Log GDP per capita		Log ANEU (in %)	
	Trend, intercept	Intercept	Trend, intercept	Intercept	Trend, intercept	Intercept
Class 1	-1.14	-0.23	-0.06	3.39	0.11	-2.36***
Class 2	-1.50*	-6.08***	-0.06	-4.28***	-6.08***	-4.32***
Class 3	0.14	-4.39***	-12.9***	0.21	-0.45	-2.25
Class 4	-0.43	-4.21***	-0.14	-2.38***	-0.01	-0.68
Class 5	0.08	2.31	-0.04	2.10	-2.99***	-4.49***
Class 6	-3.74***	-3.88***	0.22	0.02	-1.97**	0.34
Class 7	0.99	-0.33	1.78	2.11	-0.33	-3.66***

Note:
***, ** and * indicate statistical H_0 rejection at 1%, 5% or 10% level respectively

5.3 Panel cointegration test

If we have established that the order of integration is similar for all series and greater than 0, it is possible to test the long-run relationship among the variables. Several non-stationary variables with the same order of integration are cointegrated if their linear combination is stationary, i.e. I(0) processes. This means that the random shocks are just a short-run deviation from long-run equilibrium and the series did not drift away from each other across the time. The notion of cointegration was introduced by Granger in 1981, followed by Engel, Johansen and Phillips contributions. Pedroni (1995, 1997) proposed a be-directional panel conitegration test and in 1999, 2004 [38] an extension to test the relation of cointegration for more than two variables. He proposed seven statistics to test the existence of long-run equilibrium: four statistics based on within-dimension and three other based on between-dimension. Other tests exist: a bivariate conitegration test, with hypothesis of homogenous vectors of cointegration was proposed by Kao (1999) [39] or several extensions of Johansen test to apply on panel data were proposed by Larsson and all (2001), Groen and Kleibergen (2003), Breitung (2005). If the cointegration test indicates that the long-run relationship exists between the series, the error correction models (ECM) should be applied. The ECM model enables to capture the deviation from long-run equilibrium and the speed of return in case of a deviation. Otherwise the relationship between the series will be estimated by using variables in first difference.

In our paper we use Pedroni test of cointegration with null hypothesis of the absence of cointegration against the presence of cointegration. In the case of our dataset, we apply the test of cointegration on cluster 1 to determine the presence or the absence of the long-run equilibrium between per capita CO_2 emissions, per capita GDP and ANEU, as well as for clusters 1, 5 and 7 between per capita CO_2 emissions and per capita GDP. All variables are expressed in natural logarithms, so in case of variables in first dereference, the parameters represent the elasticity. As shown in **Table 5,** the null hypothesis of no cointegration is not rejected for cluster 1 and 7. This means that the long-run relationship between the variables does not exist. For cluster 5 only 1 of 7 tests (group ADF) rejects the null hypothesis at 1% in model with trend and intercept and 2 of 7 tests for the model with intercept (Panel ADF at 10% and Group ADF at 5%). It is not enough to conclude that the long-run nexus between per capita CO_2 emissions and per capita GDP exists. For all clusters and 93 countries, the long-run equilibrium does not exist between per capita GDP, per capita CO2 emissions and ANEU. Only short-run model may be estimated.

Table 5 : Results of cointegration test

Model with trend and intercept				
	Cluster 1 (a)	Cluster 1 (b)	Cluster 5 (b)	Cluster 7 (b)
Panel v-statistic	-0.11	0.61	-0.78	-0.82
Panel rho-statistic	0.46	0.45	0.57	0.31
Panel PP-statistic	-0.37	-0.48	0.18	-0.17
Panel ADF-statistic	-0.51	0.03	-0.33	-0.19
Group rho-statistic	1.38	0.8	1.62	1.51
Group PP-statistic	-1.18	-1.24	-0.39	0.96
Group ADF-statistic	0.34	-1.01	-2.3***	1.2
Model with intercept				
	Cluster 1 (a)	Cluster 1 (b)	Cluster 5 (b)	Cluster 7 (b)
Panel v-statistic	0.18	0.31	-1.08	-0.43
Panel rho-statistic	0.64	-0.74	-0.74	-0.26
Panel PP-statistic	0.25	-1.29	-1	-0.65
Panel ADF-statistic	0.28	-0.91	-1.28*	-0.49
Group rho-statistic	1.4	0.34	0.64	0.09
Group PP-statistic	0.56	-0.88	-0.78	-0.54
Group ADF-statistic	0.49	-0.78	-2.01**	-0.25

Note:
***, ** and * indicate statistical H_0 rejection at 1%, 5% or 10% level respectively
(a) test of cointegration between per capita CO2 emissions, per capita GDP and ANEU
(b) test of cointegration between per capita CO2 emissions and per capita GDP

6. Model presentation and results

The results of unit root tests show us that a part of variable is stationary for some clusters, as well as the nature of variable stationary into the cluster is not similar for all countries. To remedy the no stationary of some variables and estimate a similar model for all clusters, we apply the first difference for all variables expressed in the natural logarithm. This is equivalent to take a growth rate of variables that is easily interpretable. We suppose that the growth rate of per capita CO_2 emissions is affected by its lag growth rate of GDP per capita (in level and with lag), also the evolution of the percentage of clean energy use (ANEU). To estimate the relationship between per capita CO_2 emissions, per capita GDP and ANEU, we propose the following model:

$$gCO2_{n,t} = \alpha_n + \beta_t + \gamma_1 * gCO2_{n,t-1} + \gamma_2 * gGDP_{n,t} + \gamma_3 * gGDP_{n,t-1} + \gamma_4 * gANEU_{n,t} + \epsilon_{n,t}$$

Where $gCO2_{n,t}$ is a growth rate of per capita CO_2 emissions, $gGDP_{n,t}$ is a growth rate of per capita GDP and $gANEU_{n,t}$ is a growth rate of percentage of alternative and nuclear energy use. The elasticity between gCO2 and regressors are represented by γ. The unobserved differences between the countries into the class are captured by α (country time-invariant fixed effect) and can explain the unobserved factors such as social and local policy aspects. The time dummies β enables us to take into account temporal effects as economics and politics crises. The choice of taking time dummy variables has been motivated by an asymmetric link between growth rates of per capita emissions and per capita GDP, especially in the case of a negative growth rate of GDP. For example, during the last financial crisis, the CO_2 emissions decreased much faster than GDP. Nevertheless, the difference between growth rate of CO_2 emissions and GDP is less strong. To estimate the model, we use two econometric techniques: Within estimation with country and time fixed effects and System GMM estimator, which is a more efficient estimation in the presence of dependant variable with a lag. These two estimations will indicate good or bad model specification. In case of non significant parameter, we will remove unnecessary variables from a model.

The result of WITHIN and SYS-GMM estimation are given in **Tables 6** and **7**. We can observe that the two models give approximately same results, except for the parameter of **$gCO2_{t-1}$.** The Within model indicates a negative relation between the growth rate of per capita CO_2 emissions at time T and T-1 for 6 clusters, while SYS-GMM estimator suggests that the relation is positive, except for cluster 3. Moreover the SYS-GMM method indicates that the relation is significant only for cluster 5 at 1% and for cluster 1 at 10%. First, we can only take into account the SYS-GMM results because WITHIN estimator is severely biased. Secondly, as mentioned in section 5.1, the cluster 1 is composed of heterogeneous countries and so, it may give wrong estimations. Under these conditions, the nexus between per capita CO_2 emissions and its lag exists only for cluster 5.

The elasticity associated with the growth rate of per capita GDP is significantly higher than zero and under one, for all clusters and both WITHIN and SYS-GMM estimations. The growth rate of GDP with a lag is significantly different from zero for clusters 2 and 3 in WITHIN model, and only for cluster 3 in SYS-GMM model. The elasticity between growth rate of per capita CO_2 emissions and growth rate of ANEU is significant for 5 clusters in WITHIN model and for clusters 1, 2, 4 and 7 for SYS-GMM model. The negative sign of ANEU's parameter suggests that the increase of clean energy use enables to reduce the per capita CO_2 emissions.

Table 6 : The estimation results using WITHIN approach

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
$gCO2_{t-1}$	0.08**	-0.14***	-0.1**	-0.02	-0.14***	-0.05	-0.01
gGDP	0.72***	0.73***	0.34***	0.41***	0.42***	0.58***	0.61***
$gGDP_{t-1}$	-0.06	0.31***	0.5***	-0.11	0.11	-0.1	0.05
gANEU	-0.08**	-0.03***	-0.02***	-0.11***	-0.01	0.01*	-0.13***
R^2	0.61	0.41	0.61	0.59	0.16	0.62	0.48

Table 7: The estimation results using SYS-GMM approach

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
$gCO2_{t-1}$	0.11*	0.07	-0.09	0.12	-0.26***	0.08	0.05
gGDP	0.73***	0.68***	0.28***	0.55***	0.41***	0.52**	0.59***
$gGDP_{t-1}$	0.16	0.38	0.78***	-0.08	0.22	-0.12	0.09
gANEU	-0.073***	-0.4***	-0.027	-0.12***	-0.0009	0.018	-0.13***
AB AR(1)	-6.6***	-6.05***	-	-3.32***	-	-	-2.59***
AB AR(2)	1.18	0.24	-	0.01	-	0.11	0.38

The results of these two estimations approaches show that the relationship between per capita CO_2 emissions and its determinants is unique for each cluster, as well as the main determinants are different. We remove non-significant variables and estimate once again the model for seven clusters. We do not need any more to use SYS-GMM estimator because the growth rate of per capita CO_2 emissions does not depend on its lag, except for cluster 5. The **Table 8** summarizes the results of the final model estimation.

Table 8: Final models

	Class 1 (Within)	Class 2 (Within)	Class 3 (Within)	Class 4 (Within)	Class 5 (SYS-GMM)	Class 6 (Within)	Class 7 (Within)
$gCO2_{t-1}$	-	-	-	-	-0.23***	-	-
gGDP	0.72***	0.79***	0.34***	0.38***	0.30***	0.52***	0.62***
$gGDP_{t-1}$	-	-	0.44***	-	-	-	-
gANEU	-0.08***	-0.03***	-0.025***	-0.11***	-	-	-0.13***
R^2	0.6	0.39	0.61	0.57		0.34	0.48

The elasticity associated with the rate of per capita GDP is positive and less than 1 for all clusters. The value of the GDP elasticity is higher for countries from Western Europe, who use a lot of clean energy (cluster 2). The elasticity from cluster 1 and 3 has approximately the same value, but in the case of cluster 3 (Eastern Europe countries, no fossil fuel producers, electricity production from coal) per capita emissions depend also on the lag of per capita GDP. The smallest value of CO_2-Income elasticity is given for cluster 4 (0.34) and for cluster 5 (0.38). Though the value of elasticity is almost equal for the two groups, we cannot conclude that countries from these clusters have a same relation between CO_2 emissions and GDP. The growth rate of per capita emissions is affected by the percentage of clean energy use for cluster 4, while the dependant variable is affected by lag of CO_2 emissions for cluster 5. However, we can notice that CO_2-Income elasticity decreases according to the country's level of economic development. The elasticity for South American countries is between countries from clusters 1, 2, 3, on one side and from clusters 5, 6 on the other side.

The second determinant of CO_2 growth rate is the variation of clean energy use. The ANEU is affected negatively by the growth rate of per capita emissions for clusters 1, 2, 3, 4 and 7 and not related for clusters 5 and 6 because both groups do not have clean energy or a very little percentage. Moreover, the CO_2-ANEU elasticity depends on energy mix and percentage of clean energy. Among the 5 groups, for which CO_2-ANEU relationship exists, the value of elasticity is more important for cluster 4 (-0.11) and 7 (-0.13). If we rank countries according to their level of clean energy use, we observe that cluster 2 presents the best performance and cluster 7 the worse. But the second group has a low elasticity compared to cluster 7. This observation suggests the existence of U-inverted evolution of elasticity or decreasing returns to scale and that this relationship does not depend on the level of economic development. Countries have to make progress in clean energy with a low cost, it would reduce the per capita emissions more significantly compared to countries with high level of ANEU. However ANEU has a low effect on growth rate of emissions. For example if ANEU increases by 10%, the growth rate of per capita CO_2 emissions decreases by 1.1%. This observation suggests that clean energy have very little impact on emissions.

7. Conclusions

This paper examined the relationship between per capita CO2 emissions, per capita GDP and energy mix. We also proposed the classification methodology to solve the problem of parameters' heterogeneity. To do so, the unbalanced panel data set, covering 93 countries, and the time period from 1960 to 2008, were used. Foremost, we proposed no parametric hierarchic clustering approach, based on 17 criteria and used the Hierarchical Clustering on Principal Components (HCPC). The results of the clustering indicate the optimal partition of 93 countries into 7 groups, each with its own characteristics. The region affiliation and the level of economic development do not determine single-handedly class partition for 5 of 7 groups. Moreover, these two criteria do not determine cluster 1 (biggest countries) and cluster 4 (countries with middle level of nuclear energy and RES). This suggests that grouping the countries by region or/and by level of economic development is insufficient. The Europe Union countries are usually examined together or with two sub-panels (Western and Eastern European countries), but our results imply to split EU27 into four. The clustering results were confirmed by the test of partial homogeneity of parameters which indicates that the Panel Data Analysis is applicable for all groups, except cluster 1.

The unit root and cointegration tests showed that a long-run relationship does not exist for any clusters and the nature of stationary is different between and into the groups. The static and dynamic estimations demonstrate that the main determinants of growth rate of per capita CO2 emissions are the growth rate of per capita GDP and the percentage of alternative and nuclear energy use (ANEU), if the level is significantly different from zero. The per capita CO2 emissions depends neither on its lag nor on the lag of per capita GDP, it increases with the growth of per capita GDP and decreases with the growth of ANEU. It's found that the CO2-Income short-run elasticity is between 0.3 and 0.79 according to the groups. We also have found that the CO2-ANEU short-run elasticity depends on clean energy level. For the groups of countries with the best performance, the elasticity is equal to -0.03, while for countries with a low level of ANEU, the elasticity is -0.11 and -0.13.

The present study is innovative in two aspects. Foremost, the problem of panel heterogeneity was solved by using the Hierarchical Clustering on Principal Components and the additional dataset, which aim was to rank countries according to 17 criteria as electricity mix, region, fossils fuel endowments, size of country, contribution of electric and industrial sectors on CO_2 emissions etc. Secondly, we used the evolution of alternative and nuclear energy use as energy mix.

References

[1] G. M. Grossman and A. B. Krueger, "Economic growth and the enironment," *Quartely Journal of Economics,* vol. 110, pp. 353-377, 1995.

[2] D. Holtz-Eakin and T. Selden, "Stoking the fires? CO2 emissions and economic growth," *Journal of Public Economics,* vol. 57, pp. 85-101, 1995.

[3] R. Perman and D. I. Stern, "Evidence from panel unit root and cointegration tests that the environmental Kuznets curve does not exist," *Australian Journal of Agricultural and Resource Economics,* vol. 47, pp. 325-347, 2003.

[4] R. McKitrick and M. C. Strazicich, "Stationarity of global per capita carbon dioxide emissions: implication for global warming scenario," *Workin Paper Departement of Economics, Applachian State University.*

[5] J. E. Aldi, "Per capita carbon dioxide emissions: convergence or divergence?," *Environmental and Resource Economics,* vol. 33, pp. 533-555, 2006.

[6] S. Dinda, "Environmental Kuznets Curve hypothesis: a survey," *Ecological Economics,* vol. 49, pp. 431-455, 2004.

[7] J. B. Ang, "CO2 emissions, energy consumption and output in France," *Energy Policy,* vol. 35, pp. 4772-4778, 2007.

[8] P. K. Narayan and R. Smyth, "Energy consumption and GDP in G7 countries: new evidence from panel cointegration with structural breaks," *Energy Economics,* vol. 30, pp. 2331-2341, 2008.

[9] M. Mehara, "Energy consumption and economic growth: the case of oil exporting countries," *Energy Policy,* vol. 35, pp. 2939-2945, 2007.

[10] F. Halicioglu, "An econometric study of CO2 emissions, energy consumption and foreign trade in Turkey," *Energy Policy,* vol. 37, pp. 1156-1164, 2009.

[11] S. Dinda and D. Coondo, "Income and emission: a panel data-based cointagration analysis," *Ecological Economics,* vol. 57, pp. 167-181, 2006.

[12] C.-C. Lee and J.-D. Lee, "Income and CO2 emissions: evidence from panel unit root and cointegration tests," *Energy policy,* vol. 37, pp. 413-423, 2009.

[13] G. Marrero, "Greenhouse gases emissions, growth and energy mix in Europe," *Energy Economics,* vol. 32, pp. 1356-1363, 2010.

[14] M. S. Hossain, "Panel estimation for CO2 emissions, energy consumption, economic growth, trade opnness and urbanization of newly industrialized countries," *Energy Policy,* vol. 39, pp. 6991-6999, 2011.

[15] S. S. Sharma, "Determinants of carbon dioxide emissions: empirical evidence from 69 countries," *Applied Energy,* vol. 88, pp. 376-382, 2011.

[16] S. Niu and a. all, "Economic growth, energy conservation and emissions reduction: a comparative analysis based on panel data for 8 Asian-Pacific countries," *Energy Policy,* vol. 39, pp. 2121-2131, 2011.

[17] P. Narayan and S. Narayan, "Carbone dioxide emissions and economic growth: panel data evidence from developing countries," *Energy Policy,* vol. 38, pp. 661-666, 2010.

[18] K.-M. Wang, "Modeling the nonlinear relationship between CO2 emissions from oil and economic growth," *Economic Modelling,* vol. 29, pp. 1537-1547, 2012.

[19] H. T. Pao and C. M. Tsai, "CO2 emissions, energy consumption and economic growth in BRIC countries," *Energy Policy,* vol. 38, pp. 7850-7860, 2010.

[20] C. Lee and C. P. Chen, "Energy-income causality in OECD countries revisited: the key role of capital stock," *Energy Economics,* vol. 30, pp. 2359-2373, 2008.

[21] E. Dijkgraaf and R. J. Vollebergh, "A test of parameter homogeneity in CO2 Panel EKC estimation," *Environmental and Resource Economics,* vol. 32, pp. 229-239, 2005.

[22] N. Aslanidis, "Environmental Kuznets Curves for carbon emissions: a critical survey," *Sustainable Development Series (Fondation ENI Enrico Mattei),* 2009.

[23] M. Wagner, "The carbon Kuznets curve: a cloudy picture emitted by bad econometrics?," *Resource and Energy economics,* vol. 30, pp. 388-408, 2008.

[24] C. Burnside, "Production function regressions, returns to scale and externalities," *Journal of Monetary Economics,* vol. 37, pp. 177-201, 1996.

[25] B. H. Baltagi and J. M. Griffin, "Pooled estimators vs. their heterogeneous countreparts in context of dynamic demand for gasoline," *Journal of Econometrics,* vol. 77, pp. 303-327, 1997.

[26] B. H. Baltagi, J. M. Griffin and W. Xiong, "To pool or not to pool: homogeneous versus heterogeneous estimators applied to cigaretter demand," *Review of Economics and Statistics,* vol. 82(1), pp. 117-126, 2000.

[27] B. H. Baltagi, G. Bresson and A. Pirotte, "Comparaison of forecast performance for homogeneous, heterogeneous and shrinkage estimators. Some empirical evidence from US electricity and natural gas consumption," *Economics Letters,* vol. 76, pp. 375-382, 2002.

[28] S. Durlauf and P. Johnson, "Multiple regimes and cross-country growth behaviour," *Journal of Applied Econometrics,* vol. 10, pp. 365-384, 1995.

[29] G. Kapetanious, "Cluster Analysis of Panel Datasets using non-standard optimisation of Information Criteria," *Journal of Economic Dynamics and Control,* vol. 30(8), pp. 1389-1408, 2006.

[30] V. Sarafidis and N. Weber, "To pool or not to pool: a partially heterogeneous alternative," *MPRA paper 20814, University Library of Munich,* 2009.

[31] Hill and Smith, "à ajouter," *à ajouter,* vol. à ajouter, p. à ajouter, 1976.

[32] F. Husson, J. Josse and J. Pages, "Principal component methods - hierarchical clusteringand - partitional clustering: why would we need to choose for visualizing data?," *Technical Report - Agrocampus, Applied Mathematics Department,* 2010.

[33] C. Hsiao, "Analysis of Panel data," *Cambridge University Press,* 2003.

[34] A. Levin, C. F. Lin and C. S. J. Chu, "Unit root tests in panel data: asymptotic and finite-simple properties," *Journal of Econometrics,* vol. 108, pp. 1-24, 2002.

[35] G. S. Maddala and S. Wu, "A comparative study of unit root test with Panel Data and a new simple test," *Oxford Bulletin of Economics and Statistics,* vol. Special issue, pp. 631-652, 1999.

[36] K. S. Im, M. H. Peseran and Y. Shin, "Testing for unit roots in heterogeneous panels," *Journal of Econometrics,* vol. 115, pp. 53-74, 2003.

[37] K. Hadri, "Testing for unit roots in heterogeneous panel data," *Econometrics Journal,* vol. 3, pp. 148-161, 2000.

[38] P. Pedroni, "Panel cointegration. Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis," *Econometric Theory,* vol. 20(3), pp. 597-625, 2004.

[39] C. Kao, M. H. Chiang and B. Chen, "International R&D spillovers: an application of estimation and inference in panel cointegration," *Oxford bulletin of Economics and Statistics,* vol. 61, pp. 691-709, 1999.

Annex A

Table 9 : Country set and associate code

Country	Code	Country	Code	Country	Code	Country	Code	Country	Code
Algeria	DZA	Croatia	HRV	Iraq	IRQ	Netherlands	NLD	Sudan	SDN
Angola	AGO	Cuba	CUB	Ireland	IRL	New Zealand	NZL	Sweden	SWE
Argentina	ARG	Cyprus	CYP	Israel	ISR	Nigeria	NGA	Switzerland	CHE
Australia	AUS	Czech Republic	CZE	Italy	ITA	Norway	NOR	Syria	SYR
Austria	AUT	Denmark	DNK	Jamaica	JAM	Oman	OMN	Thailand	THA
Azerbaijan	AZE	Ecuador	ECU	Japan	JPN	Pakistan	PAK	Trinidad & Tobago	TTO
Bahrain	BHR	Egypt	EGY	Jordan	JOR	Peru	PER	Tunisia	TUN
Bangladesh	BGD	Estonia	EST	Kazakhstan	KAZ	Philippines	PHL	Turkey	TUR
Belarus	BLR	Finland	FIN	Korea, Rep.	KOR	Poland	POL	Turkmenistan	TKM
Belgium	BEL	France	FRA	Kuwait	KWT	Portugal	PRT	Ukraine	UKR
Bolivia	BOL	Germany	DEU	Latvia	LVA	Qatar	QAT	United Arab Emirates	ARE
Bosnia and Herzegovina	BIH	Greece	GRC	Libya	LBY	Romania	ROM	United Kingdom	GBR
Brazil	BRA	Guatemala	GTM	Lithuania	LTU	Russia	RUS	United States	USA
Bulgaria	BGR	Honduras	HND	Luxembourg	LUX	Saudi Arabia	SAU	Uzbekistan	UZB
Cameroon	CMR	Hungary	HUN	Malaysia	MYS	Singapore	SGP	Venezuela,	VEN
Canada	CAN	Iceland	ISL	Malta	MLT	Slovak Republic	SVK	Vietnam	VNM
Chile	CHL	India	IND	Mexico	MEX	Slovenia	SVN	Yemen	YEM
China	CHN	Indonesia	IDN	Mongolia	MNG	South Africa	ZAF		
Colombia	COL	Iran	IRN	Morocco	MAR	Spain	ESP		